DOCUMENT RESUME

ED 204 402 TM 810 416

AUTHOR McLarty, Joyce R.
TITLE Choosing Passing Scores for Tests Required for High School Graduation.
PUB DATE 15 Apr 81
NOTE 40p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Los Angeles, CA, April 14-16, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cutting Scores; *Graduation Requirements; *High Schools; *Minimum Competency Testing; Scoring; *Standards

ABSTRACT

Educators were asked to read a description of a minimum proficiency test required for high school graduation and select a cutoff score for it. The 24 descriptions used were formed as all possible combinations of a test content, a level of student performance, a penalty for failing the test, and a community response to the test. Of the 223 educators responding, 35% selected the median cutoff score of 70%. Student performance and test content were ranked the most important considerations in selecting the passing score, followed by penalty for failure and, last, community response. Analysis of variance on the cutoff score of the fully-crossed four-factor design yielded a single significant effect: the student performance by community response interaction. Post hoc contrasts indicated that a lowering of the passing score for poor student performance in the presence of community opposition was the major contributor to the effect. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED204402

Choosing Passing Scores for Tests Required for High School Graduation

Joyce R. McLarty
University of California, Los Angeles

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. McLarty

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Running head: Choosing Passing Scores

TM 810416

Abstract

Educators were asked to read a description of a minimum proficiency test required for high school graduation and select a passing (cutoff) score for it. The 24 descriptions used were formed as all possible combinations of a test content (life skills or basic skills focus), a level of student performance (high, medium, or low), a penalty for failing the test (mild or severe), and a community response to the test (supportive or opposing).

Of the 223 educators responding, 35% selected the median cutoff score of 70%. Most cutoff scores selected were multiples of five. Student failure rates of 4%, 9%, and 18% each resulted from a cutoff score selected by more than 20% of the population.

Student performance and test content were ranked the most important considerations in selecting the passing score, followed by penalty for failure and, last, community response. Community response was ranked more important for support than opposition, and test content was ranked more important for life than for basic skills. Perception of the conditions showed significant main effects of the appropriate condition except for student performance which had only a significant content by penalty interaction. Perception of the penalty had (in addition to the main effect of penalty condition), 2 significant 3-way and 2-two way interactions, and a second significant main effect.

Analysis of variance on the cutoff score of the fully-crossed four-factor design yielded a single significant effect; the student performance by community response interaction. For community support, a lower passing score was set for medium than for high or low student performance. For community opposition, a higher passing score was set for medium than for high or low student performance. Post hoc contrasts indicated that a lowering of the passing score for poor student performance in the presence of community opposition was the major contributor to the effect.

Choosing Passing Scores for Tests Required for High School Graduation

Most educators find selecting a passing score for a test challenging. Some opt out of the dilemma by designating a fixed percentage of items ("70% is passing") or a fixed percentage of students ("lowest 10% fail") as an absolute standard. The rest maneuver toward an acceptable standard through a maze of educational, ethical, and humanitarian considerations. The more serious the consequences to student, school, and community, the tougher the decision.

When a passing score determines whether or not the student will graduate from high school, the difficulties associated with selecting that score escalate. Many students are touched by the decision, and the consequences for test failure are among the most serious associated with any test score. The range and complexity of relevant considerations also expands to include financial and political constraints which are both real and powerful. Embodied in a single cutoff score, the minimum competency decision system must interface acceptably with other educational and social decision systems.

Methodologists offer a single, basic tool for selecting cutoff scores: estimation. Guided by prediction equations,

they estimate with ever-increasing accuracy a vast collection of pre-dichotomized criterion measures. These measures may be defined in terms of the expected item-level performance of hypothetical students (Angoff, 1971; Bernknopf, Curry, and Bashaw, 1979; Ebel, 1972; Nedeklsky, 1954), or by the actual item-level performance of students (Skakun and Kling, 1980). They may be defined in terms of the actual test performance of one or more groups of students (Berk, 1976; Crehan, 1974). They may be defined as the expected performance on an actual or theoretical domain of items sampled by the test, or with respect to the latent trait underlying it (Emrick, 1971; Dayton and Macready, 1976; Kriewall, 1969; Millman, 1972; Steinheiser et al., 1978). They may even be defined in terms of expert judgment as to what a high school graduate "should know" (Jaeger, 1980). Whether the criterion measure is directly observed and explicitly acknowledged or is implicit in the identification of exemplar students, the basic approach remains the same. A criterion measure is defined and dichotomized in some way. The methodologist then provides the "correct" cutoff score: the one which minimizes some function of classification error. Adjustments can be made in the equations for the relative seriousness of false-positive to false-negative errors.

Unfortunately for educators, appropriate pre-dichotomized criterion measures of minimum competence (and most other

educational outcomes) are in short supply. So are ready-made ratios of the seriousness of false-positive to false-negative errors. There are other problems as well. Even when ratios of the seriousness of types of classification errors are available, it is questionable whether they can adequately represent such diverse concerns as the financial impact of the remediation program on the school budget and the psychological trauma accruing to the student denied a high school diploma. Then, too, when the test and criterion designations disagree on the classification of a student, estimation approaches to standard-setting automatically brand the test-based classification "wrong". This assumption that the criterion measure is, by definition, more accurate than the test may be problematic, especially if the test-based decision is designed to replace, for example, a judgment-based criterion.

Existing standard-setting methodology, by assuming that the task is to estimate some external standard as accurately as possible, discourages consideration of some very important matters. School districts can rarely afford to ignore the financial, personal, political, and educational impacts imposed by the cutoff score on a test required for high school graduation.

This study explores the relationships among some commonly encountered circumstances surrounding a test required for high school graduation and between them and the cutoff score selected for that test. The circumstances include: (a) the nature of the test content (basic skills or life skills focus); (b) the general level of student performance on the test; (c) the severity of the penalty for failing the test, and (d) the degree of community support for the minimum competency assessment program. In addition to providing some baseline information on the effects of these specific conditions on the choice of a cutoff score for the test, this study aims to: (a) develop methodology for studying the process of selecting a cutoff score; (b) generate models of the standard-setting process; (c) identify areas of inquiry which are deserving of further study, and (d) sensitize educators to the information needs of standard-setters.

Method

Subjects

Subjects were 223 educators in the greater Los Angeles area. Most (161) hold certificated positions in the Simi Valley Unified School District where all of the four junior high and two senior high schools participated in the study. Eleven of the subjects were certificated educators at

Moorepark High School. Simi Valley and neighboring Moorepark are located in Ventura County, about 45 miles north-west of Los Angeles. They serve a middle to lower class white and hispanic population with small proportions of black and oriental children. There are no major urban centers in the area.

Of the remaining subjects, 41 were student teachers enrolled in classes at U.C.L.A. and 10 were community college teachers (6 at Santa Monica and 4 at Los Angeles community colleges).

Participation was voluntary except for the student teachers. No systematic data collection was done at the community colleges, so no participation rates are available for them. Participation rates averaged 45% for the junior high schools and 32% for the senior high schools. No differences in participation rates based on the pre-assigned experimental conditions were detected.

Most of the subjects were teachers (194). The rest were administrators, counselors, and other certificated personnel. Median tenure for the group was 10.5 years, with a range from 0 years (student teachers) to 35 years of experience. Forty-two percent of the subjects had taught reading at some time, and 65% reported having set a passing score before.

More than half of the subjects (121) were parents, and most (90) had children attending school. Just over a third of the subjects (78) were themselves students. Nearly ninety percent of the subjects reported that they favored proficiency testing in general, and 70% indicated that they favored it strongly.

Materials

Descriptive material was prepared for each of the four factors to be examined: (a) test content; (b) level of student performance; (c) penalty for failure, and (d) community response to the minimum competency program. Two versions were prepared for each factor except for student performance level where three versions were used. Test content was represented by a description of the test development strategy and by a sample item. The two versions both proported to be of reading tests. One reflected a basic skills approach (items related to the reading curriculum), the other a life skills approach (items related to post-high school reading tasks).

Student performance on the test was represented by an artificial but realistic score distribution presented in two forms: a frequency distribution and a table of the student failure rates which would result from each score being used as the cutoff. The three versions of student performance were created by placing the distribution at three levels on

the scale of item scores. In the hig performance condition, the top-scoring students are depicted as having reached the maximum possible score of 100 items correct. In the medium performance condition they have reached a score of 95, and in the lowest performance condition, they have reached a score of 90. Tables of failure rates were adjusted to accurately reflect the number and percent failed for each distribution placement. Thus, the same cutoff score failed successively larger proportions of students on the high, medium, and low distribution. Conversely, it was necessary to lower the cutoff score to maintain the same failure rate on them.

Penalty for failure descriptions included information on whether a diploma was awarded, whether post-high school employment would be available to those who failed the test, and what arrangements had been made for remediation. This confounded psychological, financial, and educational penalties to the student. The two versions depicted a mild penalty (diploma awarded, employment available, easy opportunities for remediation) and a severe penalty (no diploma awarded, difficulty getting employment, and few opportunities for remediation).

The community response condition described community involvement in the minimum competency assessment program. In one version the community was described as having been

actively involved in developing the program and being likely to support whatever decision was made on the cutoff score. In the other version, the community was described as having been largely excluded from the development of the minimum competency assessment program and now responding as distrusting of if not actively hostile to it. In the condition, the community was reported to be likely to challenge any decision which was made on the cutoff score.

The descriptions were collated into 24 different forms, the result of crossing two versions of test content with three of student performance, two of penalty for failure, and two of community response. Within each of the 24 forms, presentation of the four conditions was randomly ordered. This resulted in a total of 576 unique forms (24 sets of 4 conditions times 24 possible orders of the conditions within the sets). In each form, the test was described as being a measure of reading comprehension, and the sample item asked for the main idea of a paragraph. Each form was provided with an introductory letter, introducing the study and the task, and with the same set of questions. Subjects were asked to select a cutoff score for the test as described, to provide biographic data, to rank order the four conditions according to which most influenced their choice of a passing score, and to indicate on a Likert scale how they perceived each of the conditions presented.

All of the material included in the 24 forms is presented in the Appendix. All pages were cut in half and assembled into booklets 8½ by 5½ inches containing: (a) a two-page introduction; (b) six pages of conditions (one page each for penalty and community response, two pages for student performance (table followed by frequency distribution) and test content (description followed by sample item), and (c) five pages of questions (cutoff score, condition importance, condition perception, biographics, and comments, in that order).

Procedure

Assignment of subjects to experimental conditions was accomplished by randomly ordering the forms and then distributing them sequentially. With permission from the appropriate administrators, forms intended for certificated school personnel were placed in their mailboxes at the schools. A collection box was placed in a designated location at each school. Forms were distributed on Monday, January 19, 1981 and collected on Friday, January 23, 1981. Reminder notices to post daily were supplied to each school, but no check was made on their use.

Student teachers completed the forms during regularly scheduled meetings of their classes. The forms were distributed and collected by the principal investigator who then

discussed the study with each of the three classes participating. Community college teachers received the forms from a colleague, Dr. June Millet. Some returned the forms to Dr. Millet, others mailed them to the principal investigator.

All data was coded and entered into the computer by the principal investigator. A 5% sample of cases was checked for accuracy in coding and entering data, and no errors were found.

Results

Cutoff Score

The median cutoff score selected by the subjects was 70% of the items correct: 35% of the subjects selected that score. An additional 19% of the subjects selected a cutoff score of 75; 10% selected a cutoff of 60; and 9% selected a cutoff of 65. Although it was possible to select a cutoff which was not a multiple of 5, few did so.

The median student failure rate was 9%, with 23% of the subjects selecting a cutoff score which resulted in that failure rate. This was followed very closely by a 4% failure rate (selected by 22% of the subjects) and an 18% rate (selected by 21% of the subjects). Failure rates ranged from 0% to 99% with about 5% of the subjects choosing to fail

more than two-thirds of the students, and a similar percent choosing to fail less than one percent.

Analysis of variance of the fully-crossed four factor experimental design, all factors considered fixed, gave one significant effect, an interaction between the student performance distribution and the community response (see Table 1 for the ANOVA results and Figure 1 for the interaction graph). With community support, the cutoff score was set highest for

---

Insert Table 1 about here

---

the highest student performance distribution, lower for the middle distribution, and again higher for the lowest distribution. With community opposition, the pattern was reversed: a medium cutoff was set for the highest student performance distribution, a higher cutoff for the middle distribution, and a considerably lower cutoff for the lowest student performance level.

Insert Figure 1 about here

Similar failure rates were obtained for the middle student performance condition regardless of the community's response (15% for support; 18% for opposition). The failure rate was somewhat higher for support than for opposition for

the highest performance distribution (10% and 6% respectively) and much higher for support than for opposition for the lowest performance condition (30% and 16% respectively).

Condition Perception

Preliminary analysis of subjects' perceptions of the conditions presented focused on whether the conditions had been perceived as intended. Each of the four perceptions variables was analyzed with respect to its paired condition (perception of test content was analyzed, for example, by the content presented). Significant correct perception of all conditions except student performance level was noted.

Because student performance level was part of the significant effect on the cutoff score, it appeared that regardless of the stated perception it had somehow been perceived. This hypothesis was examined by analyzing each of the four perception variables as an outcome of the original four-factor design. Community response and test content showed only significant main effects of their respective conditions. Perception of the level of student performance also had a single significant effect: a content by penalty interaction (see Figure 2). Perception of the penalty for

---

Insert Figure 2 about here

---

failure had six significant effects: two three-way interactions (student performance distribution by test content by community response and student performance distribution by test content by penalty for failure), two two-way interactions (student performance distribution by community response and student performance distribution by penalty for failure), and two main effects (test content and penalty for failure). The two three-way interactions are presented in Figures 3 and 4.

Insert Figures 3 and 4 about here

Condition Importance

The subjects ranked the four conditions with respect to the importance of each in the selection of a cutoff score. Student performance and test content both received median ranks of two, with most subjects ranking one first and the other second. Penalty for failure was ranked third and community response to the testing program fourth. Subjects had the option of indicating that a condition was irrelevant to their choice of a cutoff score, and that was considered a rank of "5". More than 30% of the sample so ranked the importance of community response, 23% so ranked the penalty

for failure, 13% so ranked student performance on the test, and 11% so ranked test content.

The importance variables were also reanalyzed as outcomes of the original four-factor design. No significant effects were found for the importance of student performance or of penalty for failure, but community response importance and test content importance both had main effects of their respective conditions. Community response was ranked of greater importance when the community was supportive than when it opposed the testing program. Test content was ranked of greater importance when the test was one of life-skills than when it was one of basic skills.

Discussion

Although the data analysis is incomplete, preliminary results are provocative. The number preference for integers evenly divisible by five was overwhelming. Even allowing for the possibility that this effect was exacerbated by the format of the table of failure rates (see Appendix), it is an influence which must be anticipated in future studies. It had the effect of transforming relatively continuous variables (the cutoff score and, especially, the failure rate) into substantially categorical ones.

The interactive effect of student performance distribution and community response on the cutoff score appears to

be an attempt on the part of the subjects to "hold the line" on the failure rate when the students have performed poorly and the community is hostile. The mean cutoff score was between 70 and 73 for the remaining five combinations of community support and student performance distribution, but was lowered to about 67 for poor performance with an opposing community response. This kept the failure rate below 20%. Had the cutoff score remained in the above 70% range, the failure rate would have exceeded that level (the failure rate associated with poor student performance and community support was actually 30%).

As interesting as the actual interactive effect was the fact that community response was ranked of the lowest importance of any of the conditions presented (more than 30% of the subjects considered it irrelevant to the choice of a cutoff score), and the level of student performance was not, according to the perception measure, perceived at all. About 80% of the subjects indicated that the students had performed well on the test.

The relationship between the condition presented, its perception, and its importance to the subject, and the effect of that relationship on the cutoff score selected, remain to be explored. While the student performance by community response effect accounted for about 5% of the variance in

the cutoff score, a trial regression model using vectors for the conditions, perceptions (correct or incorrect), importance, and two-way interactions between each condition and its perception and each condition and its importance accounted for more than 20%. The preliminary ANOVA results suggest the possibility of cross-condition effects as well.

Biographic data also remains to be investigated. The preliminary analysis indicated that there was no bias in the return rate by the biographic information collected. It also ascertained that the student performance by community response interactive effect obtained in all of those subgroups with sufficient data to permit separate analysis (teachers, parents, subjects with less than 10 years tenure, subjects with more than 10 years tenure, subjects who had taught reading, and subjects who had set a passing score before).

Much model-building work remains to be done for this data. When it is complete, replications will be in order. Although teachers were selected because the task of choosing a cutoff score would be familiar to them and because a sufficient number of them were available to participate, a study with administrators and other decision-makers as subjects would be useful.

Meanwhile, the study methodology appears (within the confines of a simulation study) to work adequately. The

results suggest that standard-setters do use information about student performance and community response in their choice of a passing score, and it might be appropriate for educators to provide such information to them whenever possible.

References

Angoff, W. H. Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational measurement, Washington, D. C.: American Council on Education, 1971.

Berk, R. A. Determination of optinam cutting scores in criterion-referenced measurement. The Journal of Experimental Education, 1976, 45, 4-9.

Bernknopf, S., Durry, A., and Bashaw, W. L. A defensible model for determining a minimal cut-off score for criterion-referenced tests. Paper presented at the National Council on Measurement in Education Meeting, San Francisco, CA., April, 1979. (ERIC Document Reproduction Service No. ED 177194).

Crehan, D. D., Item analysis for teacher-made mastery tests. Journal of Educational Measurement, 1974, 4, 255-262.

Dayton, C. M. & Macready, G. B. A probabalistic model for validation of behavioral hierarchies, Psychometrika, 1976, 41, 189-204.

Ebel, R. L. Essentials of educational measurement. New Jersey: Prentice Hall, Inc., 1972.

Emrick, J. A. An evaluation model for mastery testing. Journal of Educational Measurement, 1971, 8, 321-326.

Jaeger, Richard M., Cole, Judith, Irwin, D. Michelle & Pratto, David J. An iterative structured judgment process for setting passing scores on competency tests, applied to the North Carolina High School Competency Tests in Reading and Mathematics. Greensboro, North Carolina: Center for Educational Research and Evaluation, (May) 1980.

Kriewall, T. E. Aspects and applications of criterion-referenced tests. Downers Grove, Illinois: Institute for Educational Research, 1972.

Millman, J. Determining test length: passing scores and test lengths for objectives-based tests. Los Angeles: Instructional Objectives Exchange, 1972.

Nedelsky, L. Absolute grading standards for objective tests. Educational and Psychological Measurement, 1954, 14, 3-19.

Skakun, E. N. and Kling, S. Comparability of Methods for Setting Standards. Journal of Educational Measurement, 1980, 17, 229-235.

Steinheiser, F. H., Jr.; Epstein, K. I.; Mirabella, A.; & Macready, G. Criterion-referenced testing: a critical analysis of selected models. Technical paper #306. U. S. Army Research Institute for the Behavioral and Social Sciences, 1978.

Table 1

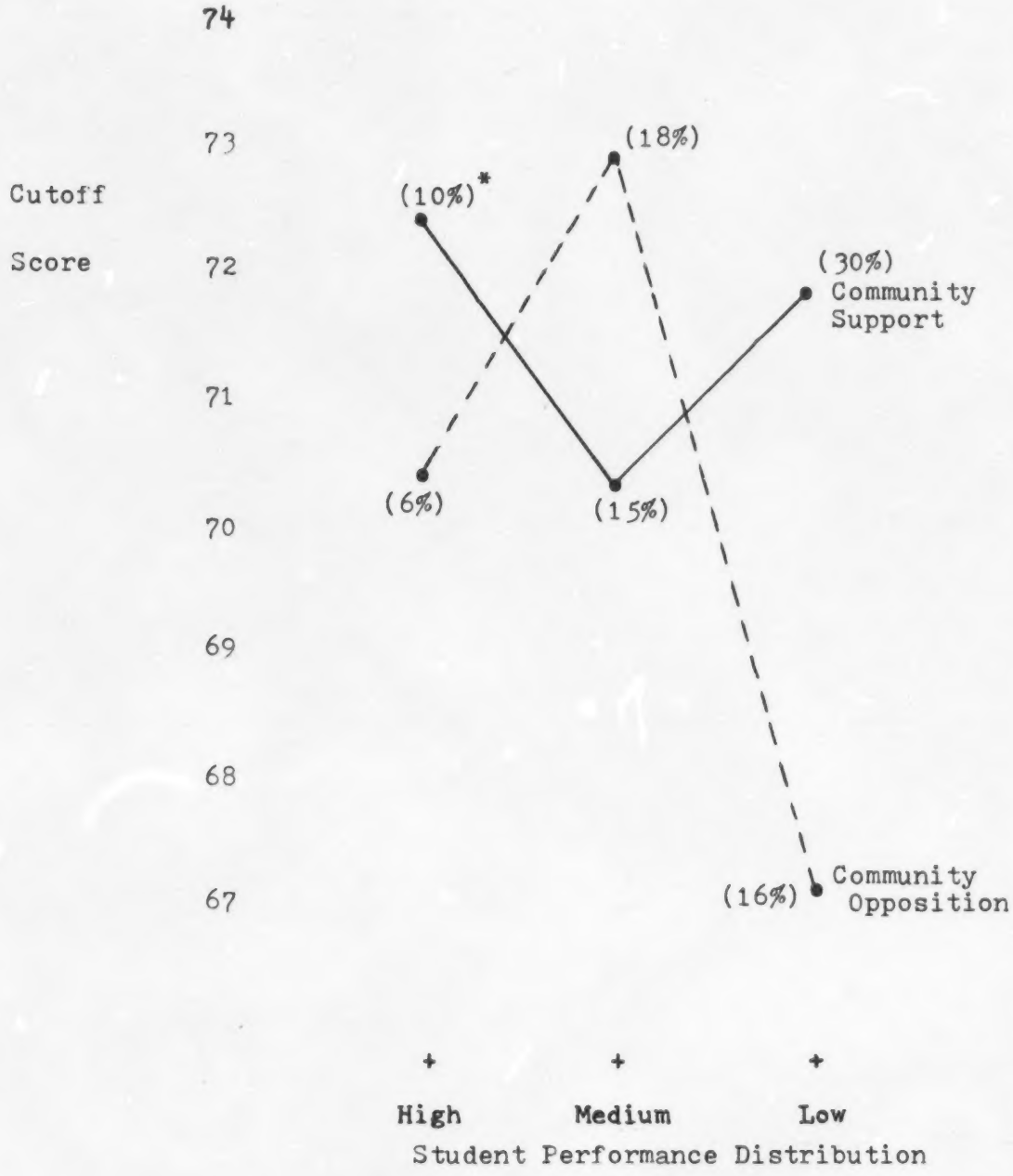
Analysis of Variance of Conditions on Cutoff Score

Source	Sum of Squares	Degrees of Freedom	Mean Square	F	Tail Probability
MEAN	980794.60317	1	980794.60317	19740.28	0.0000
DISTCOND	144.95496	2	72.47748	1.46	0.2352
CONTCOND	75.02382	1	75.02382	1.51	0.2207
PNLTCOND	15.31743	1	15.31743	0.31	0.5794
RSPNCOND	70.50367	1	70.50367	1.42	0.2351
DC	22.49198	2	11.24599	0.23	0.7977
DP	36.74061	2	18.37030	0.37	0.6914
CP	70.46837	1	70.46837	1.42	0.2352
DR	386.40819	2	193.20410	3.89	0.0222*
CR	40.65440	1	40.65440	0.82	0.3669
PR	17.80854	1	17.80854	0.36	0.5501
DCP	27.50443	2	13.75222	0.28	0.7585
DCR	15.13004	2	7.56502	0.15	0.8589
DPR	15.06530	2	7.53265	0.15	0.8594
CPR	10.44891	1	10.44891	0.21	0.6471
DCPR	24.06171	2	12.03086	0.24	0.7852
ERROR	9241.39639	186	49.68493		

*p<.05

Figure 1

Interaction of Community Response and Student Performance Distribution on the Cutoff Score



* percent of students failing the test given this cutoff score and this distribution

Figure 2

Interaction of Test Content and Penalty for Failure on Perception of Student Performance Distribution

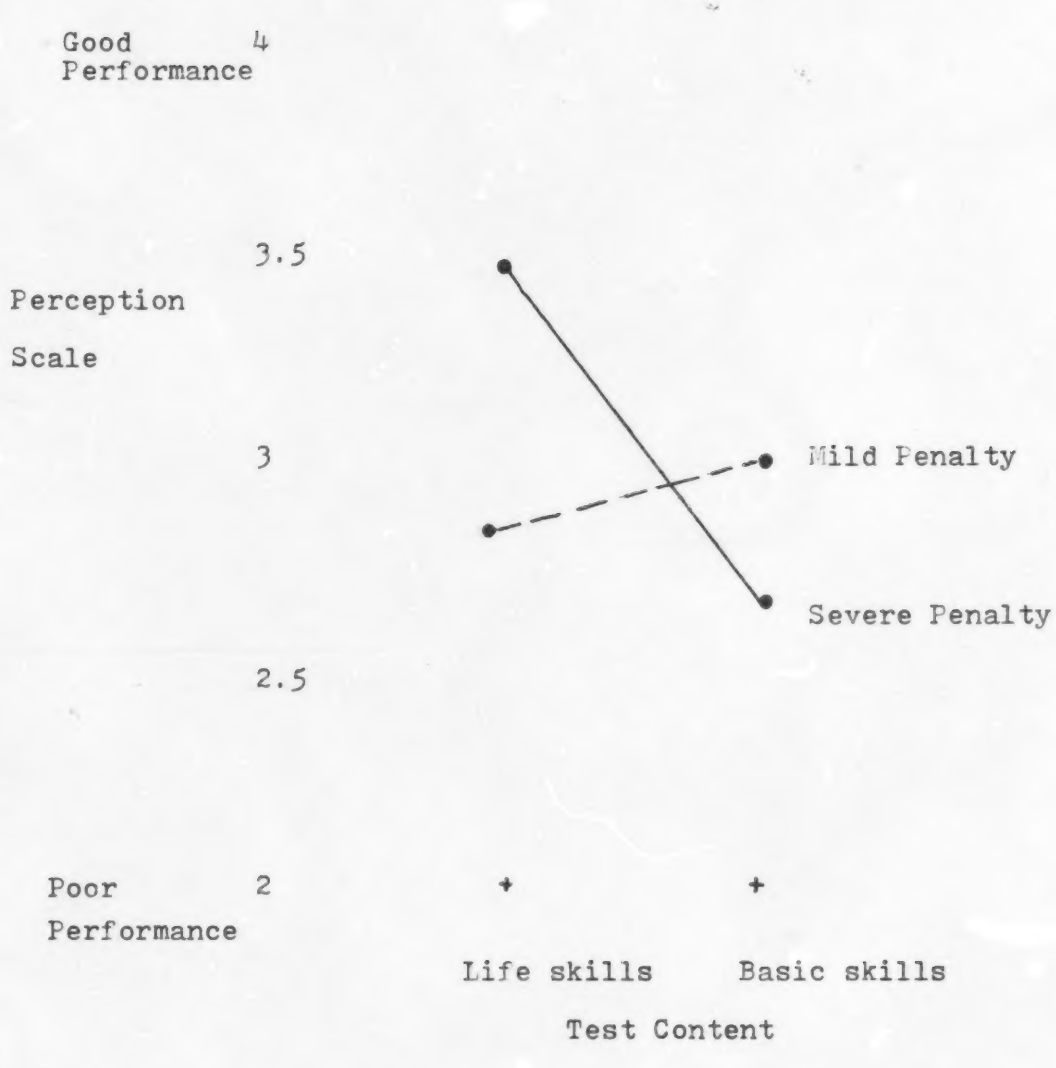


Figure 3

Three-way interaction of Student Performance Distribution, Test Content, and Penalty for Failure on Perception of the Penalty

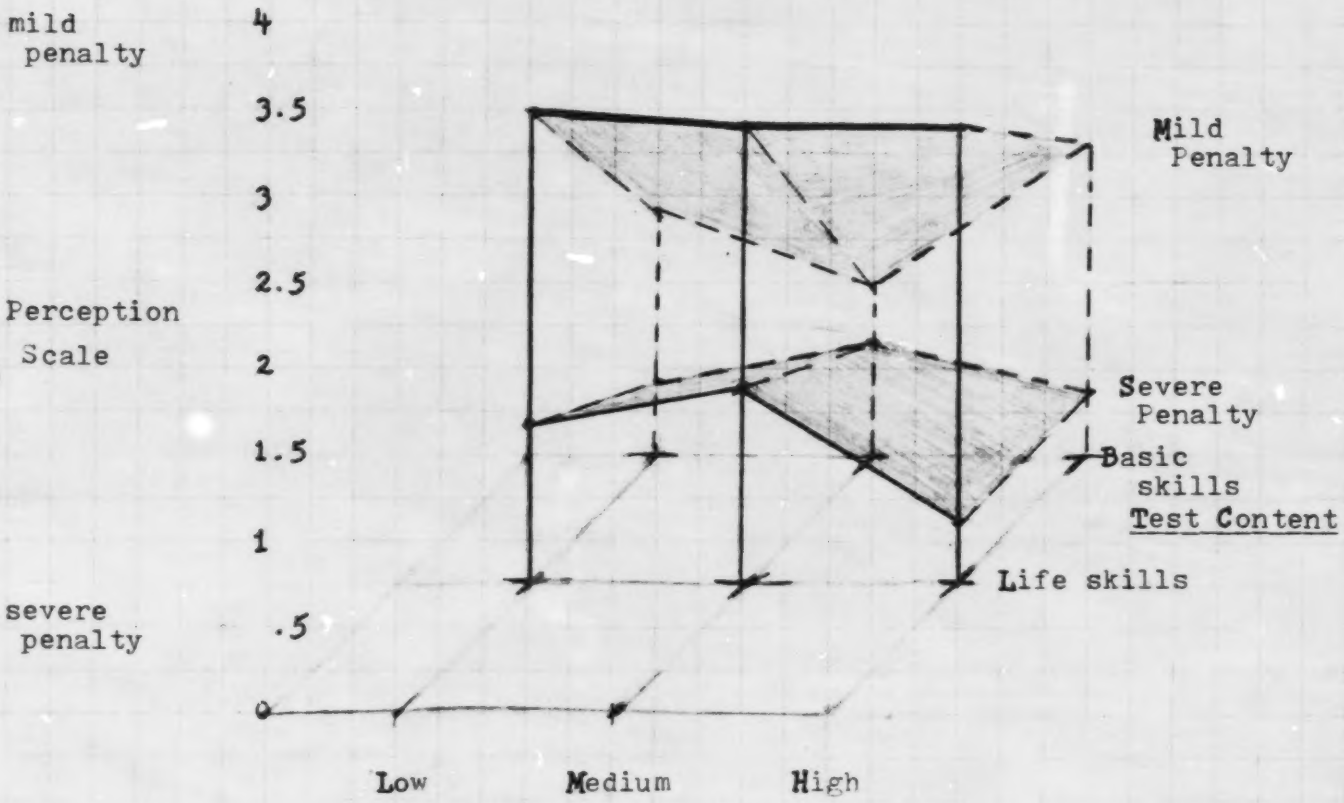
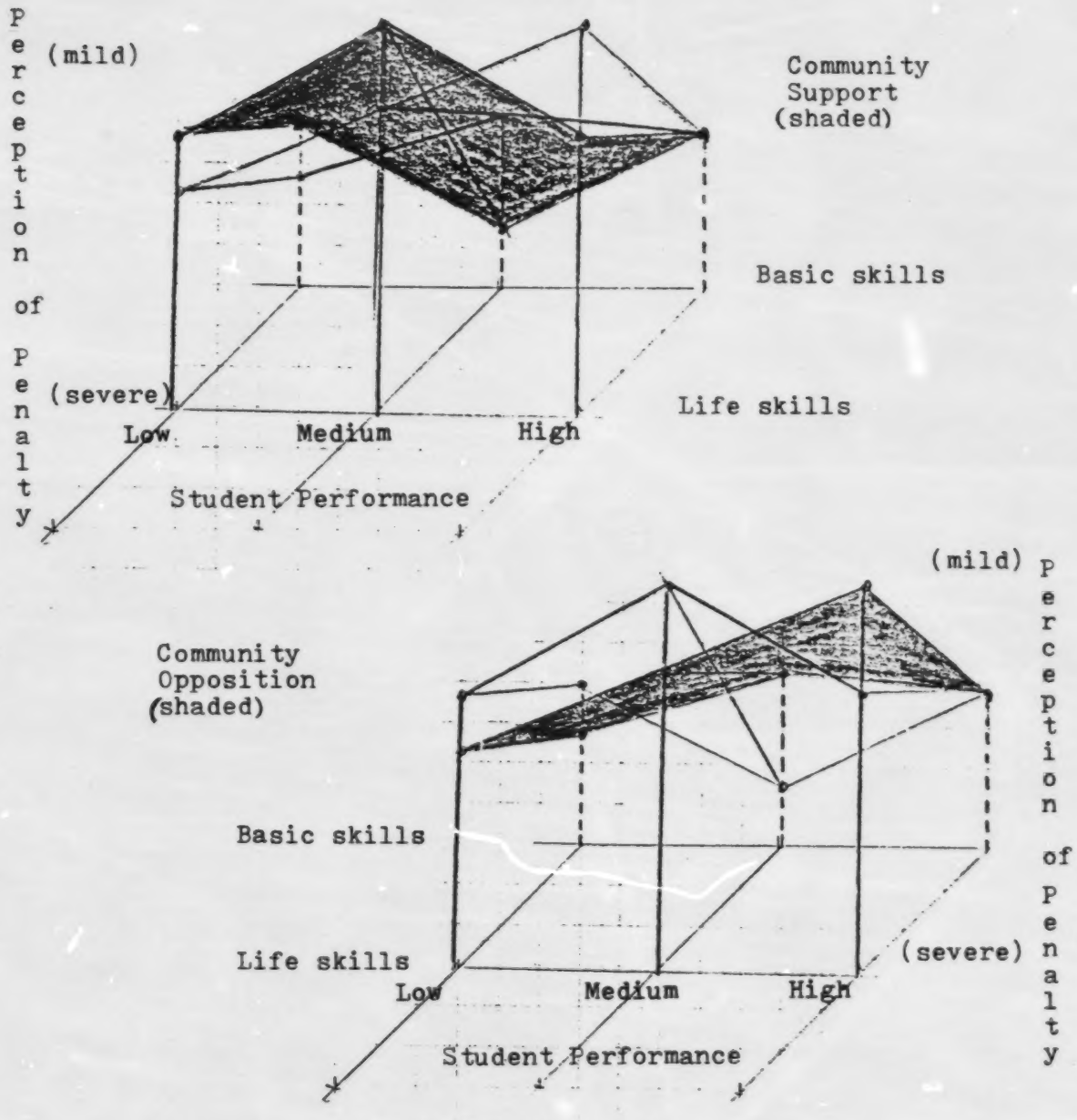


Figure 4

Three-way interaction of Student Performance Distribution, Test Content, and Community Response on Perception of the Penalty



Note: These two figures are identical except for the shading.

Appendix

Contents

Cover letter to Educators	(2 half-pages)
Test content descriptions	(2 versions)
Sample test items	(2 versions)
High Student Performance Distribution	(frequency distribution and table)
Medium Student Performance Distribution	(frequency distribution and table)
Low Student Performance Distribution	(frequency distribution and table)
Penalty for Failure descriptions	(2 versions)
Community Response descriptions	(2 versions)
Passing score chosen, favoring proficiency testing and importance of conditions	(2 half pages)
Perception of conditions, background information	(2 half pages)
Comments space	(one half page)

Dear Educator,

California now requires students to pass a minimum proficiency test in order to receive a high school diploma. As part of a research project being carried out at UCLA, we are investigating the effect of circumstances on the passing score set for such tests. Results of this investigation will be made available to assist educators who must set passing scores for proficiency tests.

Please assist us by considering the four types of circumstantial information provided and selecting a passing score for the test described. The information includes: (1) test content,

(2) student performance on the test, (3) penalty for failing the test, and (4) community response to the proficiency testing program.

This exercise takes only a few minutes, and is to be completed anonymously. Results of the research will be provided to you via your school.

Please return your completed booklet to:

_____

by: _____.

Thank you for your help.

Project on Proficiency Standards
Graduate School of Education
UCLA

Test Content - 1

This is a "life-skills" test. This means that TEST ITEMS ARE NOT NECESSARILY RELATED TO THE READING CURRICULUM in the district. Items represent the reading requirements of jobs commonly held by high school graduates (typist, auto mechanic, etc.) and of community activities (shopping, looking for a job, voting, paying taxes, etc.). Each item has been reviewed to assure that it tests an important skill needed by adults.

C1

Test Content - 1

This is a "basic-skills" test. This means that TEST ITEMS ARE CLOSELY RELATED TO THE READING CURRICULUM in the district. A detailed definition of reading comprehension was developed from the curriculum and items were created to match that definition. Each item has been reviewed to assure that it tests reading comprehension skills which are taught in the district.

C2

Test Content - 2

A typical test item:

Your water meter can detect a leak in your water system. If you have a meter, it is probably located in a small, recessed concrete vault near the street. Turn off all the faucets in and around your home, then check the meter reading. Wait thirty minutes then check the meter again. If the reading has changed, you have a leak.

The main idea of this passage is:

A. Water meters are in concrete vaults near the street
B. You can read your own water meter
C. Leaks can waste a lot of water
* D. You can use your water meter to detect leaks

C1

Test Content - 2

A typical test item:

Most seagulls don't bother to learn more than the simple facts of flight - how to get from shore to food and back again. For most gulls it is not flying that matters, but eating. For this gull, though, it was not eating that mattered, but flight. More than anything else, Jonathan Livingston Seagull loved to fly.

The main idea of this passage is:

A. Most gulls like to eat more than they like to fly
B. Most seagulls don't bother to learn very much
C. Jonathan Livingston Seagull loved to fly
* D. Jonathan Livingston Seagull was not like other seagulls

C2

Student Performance - 1

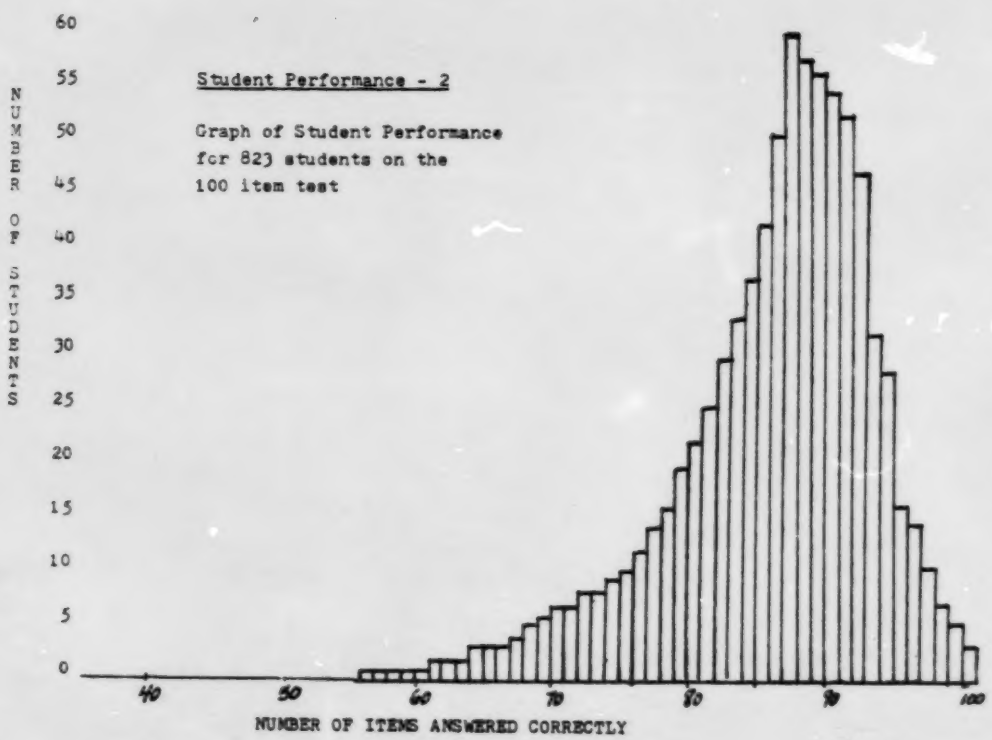
This table gives the number and percent of STUDENTS WHO WILL FAIL the test at each passing score. For example, if the passing score were set at 70 items correct (70% of the 100 items on the test), then 35 students (4% of the students who took the test) would fail the test.

Passing Score	Failing Number	Failing Percent	Passing Score	Failing Number	Failing Percent	Passing Score	Failing Number	Failing Percent
100	820	99%	85	291	35%	70	35	4%
99	815	99%	84	254	31%	69	29	3.5%
98	808	98%	83	221	27%	68	24	2.9%
97	798	97%	82	192	23%	67	20	2.4%
96	784	95%	81	167	20%	66	17	2.1%
95	768	93%	80	145	18%	65	14	1.7%
94	740	90%	79	126	15%	64	11	1.3%
93	708	86%	78	110	13%	63	9	1.1%
92	661	80%	77	96	12%	62	7	.9%
91	609	74%	76	84	10%	61	5	.6%
90	555	67%	75	74	9%	60	4	.5%
89	499	61%	74	65	8%	59	3	.4%
88	442	54%	73	57	7%	58	2	.2%
87	383	46%	72	49	6%	57	1	.1%
86	333	40%	71	42	5%	56 and below	0	0%

D1

Student Performance - 2

Graph of Student Performance for 823 students on the 100 item test



D1

Student Performance - 1

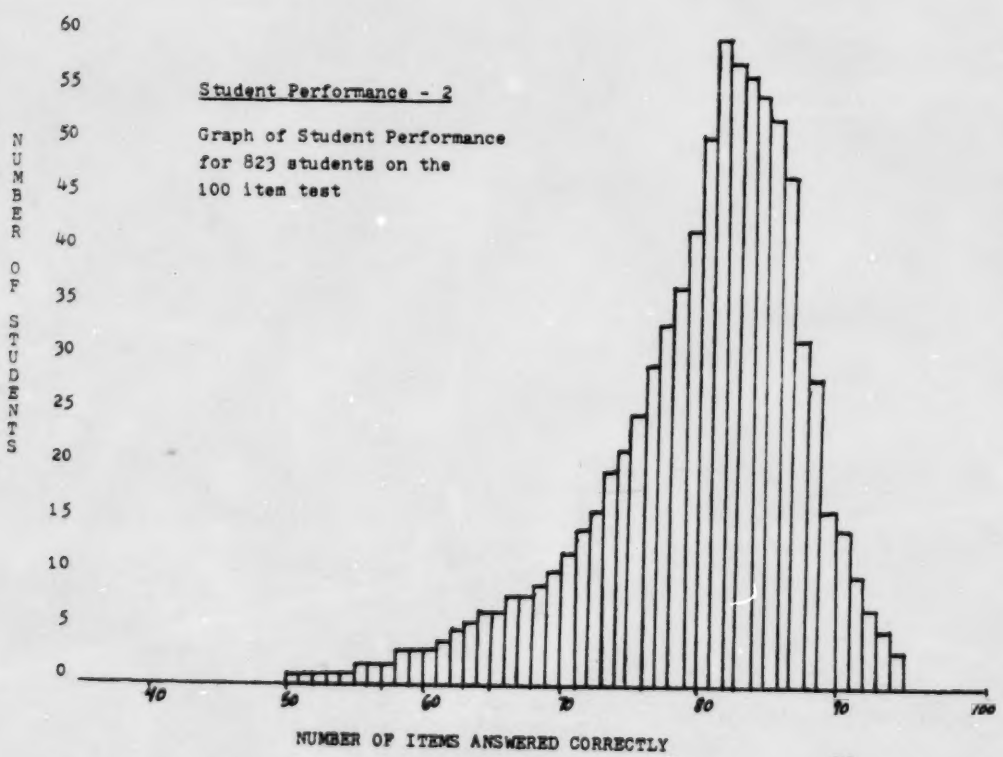
This table gives the number and percent of STUDENTS WHO WILL FAIL the test at each passing score. For example, if the passing score were set at 70 items correct (70% of the 100 items on the test), then 74 students (9% of the students who took the test) would fail the test.

Passing Score	Failing Number	Failing Percent	Passing Score	Failing Number	Failing Percent	Passing Score	Failing Number	Failing Percent
96 and above	823	100%						
95	820	99%	80	291	35%	65	35	4%
94	815	99%	79	254	31%	64	29	3.5%
93	808	98%	78	221	27%	63	24	2.9%
92	798	97%	77	192	23%	62	20	2.4%
91	784	95%	76	167	20%	61	17	2.1%
90	768	93%	75	145	18%	60	14	1.7%
89	740	90%	74	126	15%	59	11	1.3%
88	708	86%	73	110	13%	58	9	1.1%
87	661	80%	72	96	12%	57	7	.9%
86	609	74%	71	84	10%	56	5	.6%
85	555	67%	70	74	9%	55	4	.5%
84	499	61%	69	65	8%	54	3	.4%
83	442	54%	68	57	7%	53	2	.2%
82	383	46%	67	49	6%	52	1	.1%
81	333	40%	66	42	5%	51 and below	0	0%

D2

Student Performance - 2

Graph of Student Performance for 823 students on the 100 item test



D2

Student Performance - 1

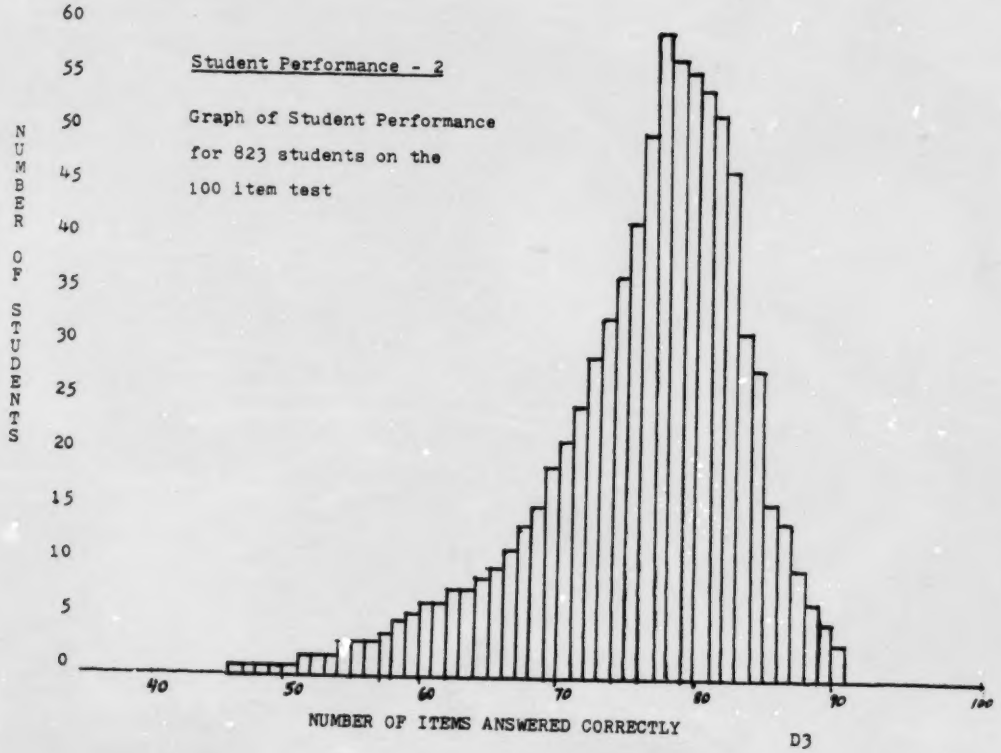
This table gives the number and percent of STUDENTS WHO WILL FAIL the test at each passing score. For example, if the passing score were set at 70 items correct (70% of the 100 items correct), then 145 students (18% of the students who took the test) would fail the test.

Passing Score	Failing Number	Failing Percent	Passing Score	Failing Number	Failing Percent	Passing Score	Failing Number	Failing Percent
91 and above	823	100%						
90	820	99%	75	291	35%	60	35	4%
89	815	99%	74	254	31%	59	29	3.5%
88	808	98%	73	221	27%	58	24	2.9%
87	798	97%	72	192	23%	57	20	2.4%
86	784	95%	71	167	20%	56	17	2.1%
85	768	93%	70	145	18%	55	14	1.7%
84	740	90%	69	126	15%	54	11	1.3%
83	708	86%	68	110	13%	53	9	1.1%
82	661	80%	67	96	12%	52	7	.9%
81	609	74%	66	84	10%	51	5	.6%
80	555	67%	65	74	9%	50	4	.5%
79	499	61%	64	65	8%	49	3	.4%
78	442	54%	63	57	7%	48	2	.2%
77	383	46%	62	49	6%	47	1	.1%
76	333	40%	61	42	5%	46 and below	0	0%

D3

Student Performance - 2

Graph of Student Performance for 823 students on the 100 item test



D3

Penalty for Failing

Students who fail the test will incur MILD PENALTIES as a result. They will receive diplomas without a proficiency certification statement. This will allow them to take jobs or enroll in educational programs open to high school graduates while receiving remedial instruction. Special remedial classes with day and evening sections are available as are tutors and counselors. Students may retake the test as many times as needed to earn their proficiency certification.

P1

Penalty for Failing

Students who fail the test will incur SEVERE PENALTIES as a result. They will be denied diplomas and will therefore be unable to take jobs or enroll in educational programs open only to high school graduates. They will have to enroll in high school for at least one more semester of course work to be eligible to retake the test with the next class. There are no provisions for special remedial instruction or individual retesting.

P2

Community Response

The community is VERY SUPPORTIVE of the proficiency testing program. Leaders of minority and majority community groups have participated in selecting the proficiencies and the types of test items to be used. Most meetings of the program staff have been open to the public, and many have included opportunities for citizens to express their concerns. These concerns have been addressed satisfactorily and the community is now ready to support any decision which is reached on the actual passing score.

R1

Community Response

The community is NOT SUPPORTIVE of the proficiency testing program. Leaders of both minority and majority community groups have expressed strong opposition to decisions reached by the program staff on the proficiencies to be tested and the types of test items to be used. The teachers' union and the PTA have expressed dissatisfaction with the tests and the implementation schedule. Any decision which is reached on the actual passing score will probably result in community protest.

R2

Please write the passing score you have chosen in this box:

Are you personally in favor of requiring students to pass a minimum proficiency test in order to graduate from high school? (mark one)

_____ yes, definitely so
_____ yes, probably so
_____ no, probably not
_____ no, definitely not

You were given four kinds of information. Please mark a "1" by the information which <u>most influenced</u> your choice of a passing score, "2" by the next most, and so on. Leave any information which did not influence your choice blank.

_____ Test Content
_____ Penalty for Failing the test
_____ Student Performance on the test
_____ Community Response to the proficiency testing program

What additional information, if any, would have helped you to select an appropriate passing score?

Please indicate how you perceived the information presented by circling SA if you strongly agree with the statement; A if you tend to agree; D if you tend to disagree; or SD if you strongly disagree.

SD	D	A	SA	The test content is closely related to the reading curriculum in the district.
SD	D	A	SA	In general, the students performed well on the test.
SD	D	A	SA	The community supports the proficiency testing program.
SD	D	A	SA	The penalty for failing the test is relatively mild.

Please provide information about your background by answering the following questions.

What is the name of your school district?_____

With what grade level(s) of students do you usually work?_____

How many years have you worked in education?_____

Are you primarily a/an ____classroom teacher?

___administrator? ___counselor? ___other?

Are you now a student? ___yes ___no

Are you a parent? ___yes ___no (have you any children still in school? ___yes ___no)

Have you ever taught reading? ___yes ___no

Have you ever set a passing score before? ___yes ___no

Comments? We'd like to know how you feel about minimum proficiency testing in general and this research in particular ...

Again, thank you for your help!